# PATTERN RECOGNITION FOR MANUFACTURING PROCESS VARIATION USING INTEGRATED STATISTICAL PROCESS CONTROL – ARTIFICIAL NEURAL NETWORK

AHMAD AZRIZAL BIN MOHD ARIFFIN

A project report submitted in fulfillment of the requirement for the award of the Degree of Master of Mechanical Engineering

Faculty of Mechanical and Manufacturing Engineering
Universiti Tun Hussein Onn Malaysia

JULY 2015

## ABSTRACT

Variation in manufacturing process is known to be a major source of poor quality products and variation control is essential in quality improvement. In bivariate cases, which involve two correlated quality variables, the traditional statistical process control (SPC) charts are known to be effective in monitoring but they are lack of diagnosis. As such, process monitoring and diagnosis is critical towards continuous quality improvement. This becomes more challenging when involving two correlated variables (bivariate), whereby selection of statistical process control (SPC) scheme becomes more critical. In this research, a scheme to address balanced monitoring and accurate diagnosis was investigated. Investigation has been focused on an integrated SPC - ANN model. This model utilizes the Exponentially Weighted Moving Average (EWMA) control chart and ANN model in two-stage monitoring and diagnosis technique. This scheme was validated in manufacturing of hard disc drive. The study focused on bivariate process for cross correlation function, $\rho = 0.3$ and 0.7 and mean shifts, $\mu = \pm 1.00$-2.00 standard deviations. The result of this study, suggested this scheme has a superior performance compared to the traditional control chart. In monitoring, it is effective in rapid detection of out of control without false alarm. In diagnosis, it is able to accurately identify for source of variation. This scheme is effective for cases variations of such loading error, offsetting tool and inconsistent pressure. Therefore, this study should be useful in minimizing the cost of waste materials and has provided a new perspective in realizing balanced monitoring and accurate diagnosis in BQC.

# ABSTRAK

Variasi dalam proses pembuatan diketahui sebagai punca utama produk berkualiti rendah dan kawalan variasi adalah penting dalam peningkatan kualiti. Dalam kes-kes yang melibatkan dua pembolehubah yang berkolerasi, carta tradisional kawalan proses statistik (SPC) diketahui berkesan dalam pemantauan tetapi ianya mempunyai kekurangan dalam aspek diagnosis. Oleh itu, pemantauan dan proses diagnosis adalah penting ke arah peningkatan kualiti yang berterusan. Hal ini menjadi lebih mencabar apabila melibatkan dua pembolehubah yang berkolerasi, di mana pemilihan skim kawalan proses statistik (SPC) menjadi lebih kritikal. Dalam penyelidikan ini, satu kajian telah dijalankan secara terperinci untuk memastikan skim yang dihasilkan dapat mengenal pasti pemantauan seimbang dan diagnosis secara tepat. Kajian ini difokuskan kepada satu model SPC bersepadu- reka bentuk rangkaian neural tiruan (ANN). Model ini menggunakan carta kawalan Purata Bergerak Pemberat Exponen (EWMA) dan model ANN dalam pemantauan dua peringkat dan teknik diagnosis. Skim ini telah disahkan dalam proses pembuatan pemacu cakera keras. Kajian tertumpu kepada proses dua pemboleh ubah, di mana nilai untuk fungsi korelasi silang, $\rho = 0.3$ dan 0.7 dan purata anjakan, $\mu = \pm$ 1.00-2.00 sisihan piawai. Hasil daripada kajian ini, skim ini telah terbukti dapat menghasilkan prestasi yang unggul berbanding dengan carta kawalan tradisional. Dalam aspek pemantauan, ianya berkesan untuk mengesan proses luar kawalan secara pantas tanpa penggera palsu. Dalam aspek diagnosis pula, ia dapat mengenal pasti dengan tepat sumber berlakunya variasi. Skim ini dapat menganalisis secara berkesan bagi kes-kes variasi seperti ralat pemasangan, mata alat tersasar dan masalah tekanan yang tidak konsisten. Oleh yang demikian, kajian ini adalah sangat berguna dalam mengurangkan kos bahan-bahan buangan dan telah memberikan perspektif baru dalam merealisasikan proses pemantauan seimbang dan diagnosis tepat dalam BQC.

# TABLE OF CONTENTS

# LIST OF FIGURES

# LIST OF TABLES

# LIST OF APPENDICES

# CHAPTER 1

# INTRODUCTION

## 1.1 Research Background

The manufacturing environment in which quality engineering is practiced is changing rapidly, with many companies facing higher demands with the introduction of new systems and new products. System transitions are becoming a more significant part of overall operations. Furthermore, there is increased pressure for quality engineering as well as other manufacturing activities to support the economic objectives and profitability of the firm. Quality engineers need more tools to cope with these changes and to meet the intense international competition .Quality control (QC) is a procedure or set of procedures intended to ensure that a manufactured product or performed service adheres to a defined set of quality criteria or meets the requirements of the client or customer. Poor quality due to process variation is known as a major issue in manufacturing processes. Manufacturing process may involve two or more correlated variables and an appropriate procedure is required to monitor these variables simultaneously. In manufacturing industries, when quality feature of a product involves two correlated variables (bivariate), an appropriate SPC (Statistical Process Control) charting scheme is necessary to monitor and diagnose these variables jointly. Specifically, process monitoring refers to the identification of process condition either in a statistically in control or out of control, whereas process diagnosis refers to the identification of the source variable(s) for out of control condition.

Statistical process control (SPC) has developed into a rich collection of tools for monitoring a system. The first control chart, proposed by Shewhart (1926), is still the one most used in industrial systems today. The general assumption under which this chart is designed is that the observational data used to construct the chart are independent and identically normally distributed, although this need not be the case (Woodall (1987). The process is declared in-control as long as the points on the chart stay within the control limits. If a point falls outside those limits, an out-of-control situation is declared, and a search for an assignable cause is initiated.

## 1.2 Problem Statement

In manufacturing industries, process variation is known to be a major source of poor quality. As such, process monitoring and diagnosis is critical towards continuous quality improvement. This becomes more challenging when involving two or more correlated variables. Process monitoring refers to the identification of process status either it is running within a statistically in-control or out-of-control condition, whereas process diagnosis refers to the identification of the source variables of out-of-control process. The traditional statistical process control (SPC) charting schemes were known to be effective in monitoring aspect. Nevertheless, they are lack of diagnosis. In recent years, the artificial neural network (ANN) based pattern recognition schemes are mainly utilize either a generalized (single) model ANN pattern recognizer and/or raw data based input representation, which resulted in limited performance.

In HDD products, motor function was to control disc movement path. Disc diameter (ID1 and ID2) as shown in Figure 1.1 is the critical features of quality that requires bivariate quality control (BQC). Die casting is a process for producing the HDD, it is suitable for forming small parts that can be found on the hard disk in order to get a smooth surface quality and dimensional consistency. After the die casting process, the next process will occur is machining process, where it will form the rough size and shifted to the actual size until the completion of machining process for forming the components on the HDD. The next machining processes used to make large dimension and a small dimension to achieve tolerance to motor installation on the HDD. Next coating process used before the installation process is done. Therefore, based on the pattern recognition of bivariate data output, (Positioning and Concentricity) this scheme will be able to monitor and diagnose process variation in mean shifts for process manufacturing of computer HDD.

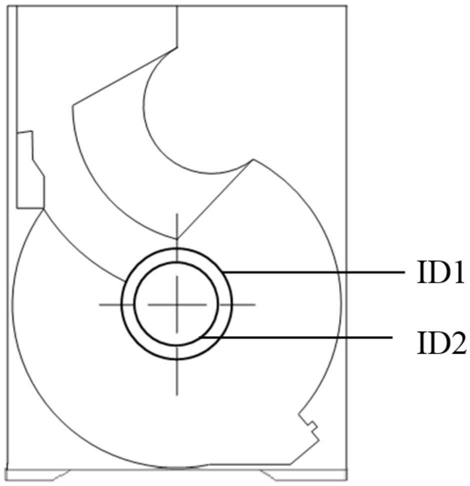



Figure 1.1: Critical features of quality that requires bivariate quality control (BQC)

In this study, an integrated ANN model that is called “Synergistic” pattern recognizer involves utilization of raw data-based and statistical features input representations. Soon, practitioners realized that the ability of the Shewhart chart to detect small changes was not as good as its ability to detect large changes. One approach to improving the sensitivity of the chart was the use of several additional rules (see, for example, Western Electric (1958). Another approach was to design complementary charts, which could be used in conjunction with the Shewhart chart, but which were better at detecting small changes. Page (1954) and developed the cumulative sum (CUSUM) chart, where past and present data are accumulated to detect small shifts in the mean. Roberts (1959) proposed the exponentially weighted moving average (EWMA) chart as another way to detect small changes. This ability comes from the fact that the EWMA statistic can be written as a moving average of the current and past observations, where the weights of the past observations fall off exponentially, as in a geometric series.

## 1.3 Project Objectives

The objectives of this project are:-

(i) To investigate the effectiveness of an Integrated SPC-ANN pattern recognition scheme for monitoring and diagnosis manufacturing process variation. Integrated SPC-ANN is used as a processing tool to assess the feasibility of processing data to monitor and diagnose process variation in mean shifts.

(ii) To evaluate the performance of the scheme in actual manufacturing process application.

## 1.4 Project Scopes

This research proposal project scope is listed as below:

(i) The bivariate process variables are dependent to each other based on linear cross correlation (ρ).

(ii) The predictable patterns of process variation are limited to sudden shifts (upward shift).

(iii) Magnitudes of variation (sudden shifts) are limited within ± 3 standard deviations based on control limits of Shewhart control chart.

(iv) Design and modeling of input data representation in training and pre-testing ANN-based model are based on Lehman (1977) model, whereas the validation tests are performed using three types of variation in process manufacturing for computer Hard Disc Drive (HDD).

# CHAPTER 2

# LITERATURE REVIEW

This chapter provided the reviews of the concept of SPC control chart monitoring and diagnosis. It were also review the investigation and development of the integrated bivariate SPC-ANN schemes combined the traditional SPC chart(s) with an ANN model. The traditional SPC chart(s) role for monitoring the existence of unnatural variation in bivariate process, whereas an ANN model roles for diagnosing the sources of variation. In that case, an ANN model is utilized only when necessary, that is, when an out-of-control signal is triggered. Inversely, the other schemes such as novelty detector ANN consist of fully ANN or fully ML-based model for monitoring and diagnosing simultaneously. In that case, an ANN model is continuously utilized, that is, for triggering out-of-control signal and then, for identifying the sources of variation. In conclusion, explanation on why the Integrated Statistical Process Control (SPC) - ANN pattern recognition is use to improve manufacturing process variation was given.

## 2.1 Introduction

In the production HDD, the production process is carried out continuously. Therefore, at regular time intervals, samples of data will be taken to investigate the variation and validate the pattern recognition on the production of HDD.

If a product is to meet or exceed customer expectations, generally it should be produced by a process that is stable or repeatable. More precisely, the process must be capable of operating with little variability around the target or nominal dimensions of the product's quality characteristics. Statistical process control (SPC) is a powerful collection of problem-solving tools useful in achieving process stability and improving capability through the reduction of variability. SPC is one of the greatest technological developments of the twentieth century because it is based on sound underlying principles, is easy to use, has significant impact, and can be applied to any process. Its seven major tools are the histogram or stem-and-leaf plot, the check sheet, the Pareto chart, the cause-and-effect diagram, the defect concentration diagram, the scatter diagram, and the control chart.

This chapter has three objectives. The first is to present the basic statistical control process (SPC) problem solving tools, called the magnificent seven, and to illustrate how these tools form a cohesive, practical framework for quality improvement. These tools form an important basic approach to both reducing variability and monitoring the performance of a process, and are widely used in both the analyze and control steps of DMAIC. The second objective is to describe the statistical basis of the Shewhart control chart. The reader will see how decisions about sample size, sampling interval, and placement of control limits affect the performance of a control chart. Other key concepts include the idea of rational subgroups, interpretation of control chart signals and patterns, and the average run length as a measure of control chart performance. The third objective is to discuss and illustrate some practical issues in the implementation of SPC.

## 2.2 Process Variation

In order to reduce variation in manufacturing processes consisting of several discrete stages it is often worthwhile to study the variation that is added at different stages, and whether that variation is transmitted downstream to subsequent stages. In particular, there may be certain stages where considerable variation originates, and other stages that filter out variation introduced upstream. Process variation is known to be a major source of poor quality (Zainal Abidin & Masood, 2012). Traditionally, statistical process control (SPC) was used to monitor and identify process variation. Advances, variation reduction efforts as such process monitoring and diagnosis should be critically applied towards quality improvements (Masood & Hassan, 2009).

Variation may be defined as any unwanted condition or as the difference between a current and a desired end-state. Both product performance and manufacturing processes exhibit variation. Wear and tear, vibration, machine breakdown, inconsistent raw material and lack of human operators' skills are the common sources of variation in manufacturing process. To manage and reduce variation, the variation must be traced back to its source. Variation occurs in all natural and man-made processes. If variation cannot be measured, it is only because the measurement systems are of insufficient precision and accuracy. Process variance reduces the capacity of the industries because processes become either under- or over-utilized. Process variance reduces the ability to detect potential problems and increases the difficulty of discovering the root cause of problems.

In any production process, regardless of how well designed or carefully maintained it is, a certain amount of inherent or natural variability will always exist. This natural variability or "background noise" is the cumulative effect of many small, essentially unavoidable causes. In the framework of statistical quality control, this natural variability is often called a "stable system of chance causes." A process that is operating with only chance causes of variation present is said to be in statistical control. In other words, the chance causes are an inherent part of the process. Other kinds of variability may occasionally be present in the output of a process. This variability in key quality characteristics usually arises from three sources: improperly adjusted or controlled machines, operator errors, or defective raw material. Such variability is generally large when compared to the background noise, and it usually represents an unacceptable level of process performance. We refer to these sources of variability that are not part of the chance cause pattern as assignable causes of variation. A process that is operating in the presence of assignable causes is said to be an out-of-control process.

These chance and assignable causes of variation are illustrated in Fig. 2.1. Until time $t_1$ the process shown in this figure is in control; that is, only chance causes of variation are present. As a result, both the mean and standard deviation of the process are at their in-control values (say, $\mu_0$ and $\sigma_0$). At time $t_1$ an assignable cause occurs. As shown in Fig. 2.1, the effect of this assignable cause is to shift the process mean to a new value $\mu_1 > \mu_0$. At time $t_2$ another assignable cause occurs, resulting in $\mu = \mu_0$, but now the process standard deviation has shifted to a larger value $\sigma_1 > \sigma_0$. At time $t_3$ there is another assignable cause present, resulting in both the process mean and standard deviation taking on out-of-control values. From time $t_1$ forward, the presence of assignable causes has resulted in an out-of-control process. Processes will often operate in the in-control state for relatively long periods of time. However, no process is truly stable forever, and, eventually, assignable causes will occur, seemingly at random, resulting in a shift to an out-of-control state where a larger proportion of the process output does not conform to requirements. For example, note from Fig. 2.1 that when the process is in control, most of the production will fall between the lower and upper specification limits (LSL and USL, respectively). When the process is out of control, a higher proportion of the process lies outside of these specifications.

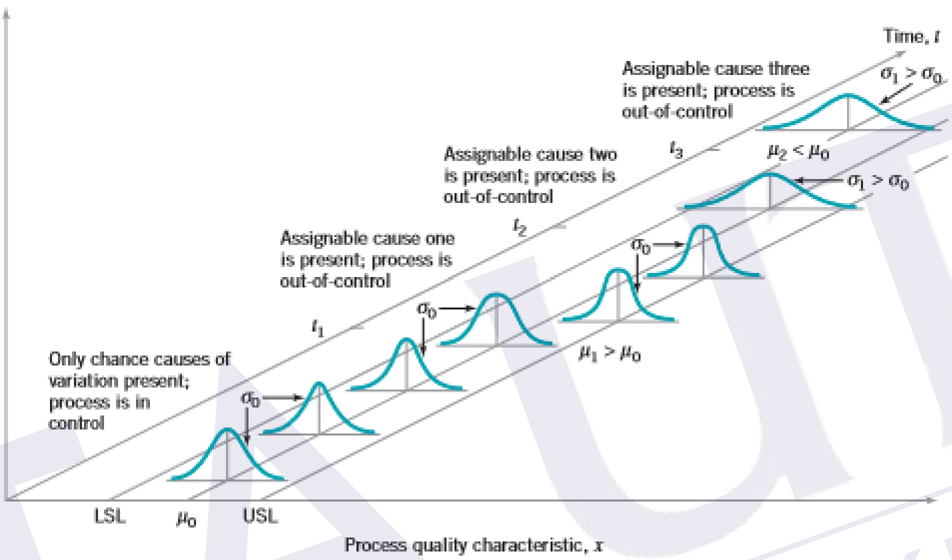



Figure 2.1: Chance and assignable causes of variation (Montgomery,Douglas, C, 2009).

A major objective of statistical process control is to quickly detect the occurrence of assignable causes of process shifts so that investigation of the process and corrective action may be undertaken before many nonconforming units are manufactured. The control chart is an on-line process-monitoring technique widely used for this purpose. Control charts may also be used to estimate the parameters of a production process, and, through this information, to determine process capability. The control chart may also provide information useful in improving the process. Finally, remember that the eventual goal of statistical process control is the elimination of variability in the process. It may not be possible to completely eliminate variability, but the control chart is an effective tool in reducing variability as much as possible.

## 2.3 Statistical Process Control (SPC)

Statistical process control (SPC) is one of the most effective tools quality management (TQM), which is used to monitor and minimize process variations. Control charts are the most widely applied SPC tools used to reveal abnormal variations of monitored measurements. Common causes are considered to be due to the inherent nature of normal process. Assignable causes are defined as abnormal shock to the processes, which should be identified and eliminated as soon as possible. When an abnormal variation is signaled by control chart, quality practitioners or engineers search for the assignable causes and take some necessary correction and adjustments to bring the out-of-control process back to the normal state. In many quality control settings, the manufacturing process may have two or more correlated quality characteristics and an appropriate approach is needed to monitor all these characteristics simultaneously. The usual practice has been to maintain a separate chart for each characteristic. However, this could result in some fault out-of-control alarms when the characteristics are highly correlated.

SPC is a technique used in a manufacturing environment to ensure quality parts are produced. Montgomery (2013) highlighted statistical process control is one of the most effective tools of total quality management whose main function is to monitor and minimize process variations. There are many ways to implement process control. Key monitoring and investigating tools included are illustrates in Fig. 2.2.

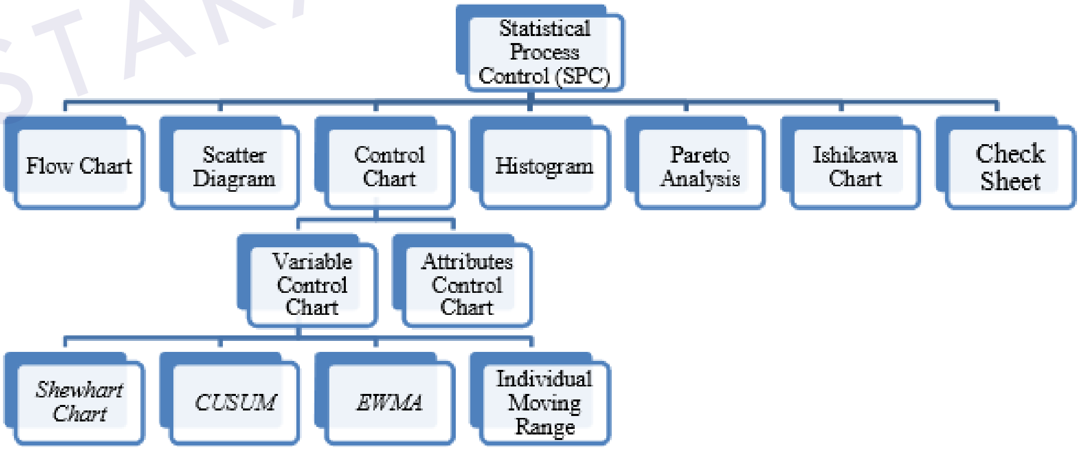



Figure 2.2: Basic statistical process control tools specification (Montgomery,Douglas, C, 2009).

A control chart is the primary tool of SPC and is basically used to monitor the process characteristics, e.g., the process mean and process variability Montgomery, 2013). The most common types of variable control charts for variables include: (1) Average and Range (X bar and R) Charts (2) Average and Standard Deviation (X and S) Charts (3) Individual and Moving Range (X and MR) Charts. Among applied tools, Shewhart control chart are the most widely applied SPC tools used to reveal abnormal variations of monitored measurements (Yu & Xi, 2009). The uses of control charts are to plot measurements of part dimensions being produced. These charts are used to alert the operator to shifts in the mean of the measurement.

In order to achieve global competitive advantage, every organization is trying to improve its product quality at each stage of the manufacturing process. Statistical process control (SPC) is one of the most effective tools of total quality management, which is used to monitor process variations and improve the quality of production. Control charts, mostly in the form of X bar chart, are widely used as aids in maintaining quality and achieving the objective of detecting trends in quality variation before defective parts/products are actually produced. In any continuous manufacturing process, variations from the established standards are mainly of two types. One is assignable cause variation, such as those due to faulty manufacturing equipment or irresponsible personnel or defective material or a broken tool. The other one is normal chance variation, resulting from the inherent non-uniformities that exist in machines or operators or materials or processes. The X bar chart usually exhibits various types of patterns [1, 2], e.g., normal (NOR), stratification (STA), systematic (SYS), increasing trend (UT), decreasing trend (DT), upward shift (US), downward shift (DS), cyclic (CYC), and mixture (MIX), as shown in Fig. 2.3.

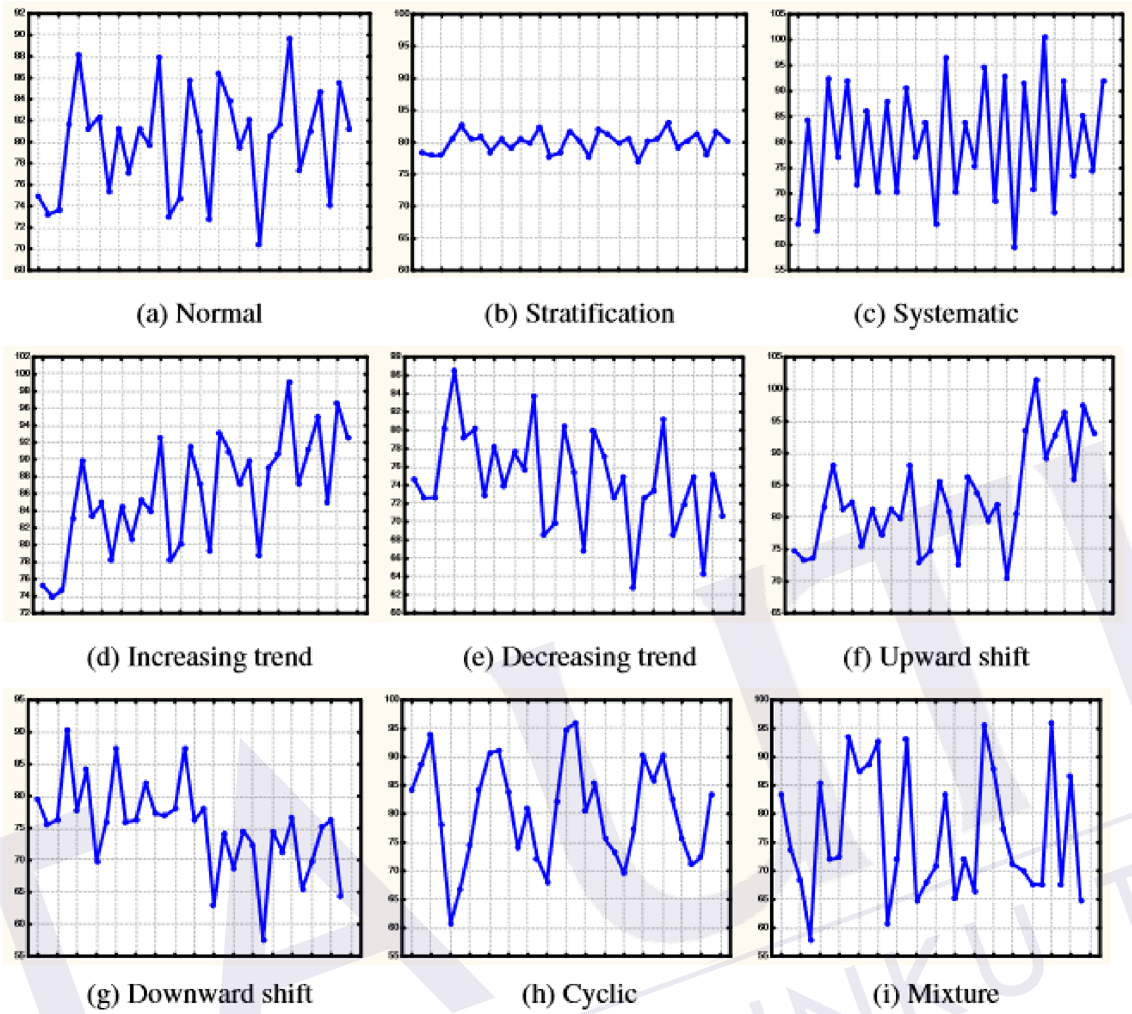



Figure 2.3: Nine control chart patterns (Montgomery,Douglas, C, 2009).

Only the normal pattern is indicative that the process is operating under random chance causes, i.e., in statistical control. The remaining patterns are unnatural and are associated with impending problems requiring pre-emptive actions. The task of control chart pattern (CCP) recognition is basically associated to accurately identify the unnatural CCPs so that prompt corrective actions can be initiated by the operators. Identification and analysis of the unnatural patterns require considerable experience and skill from the part of the quality control practitioners. However, usually, they are lacking the skill and expertise needed for interpretation of the control chart patterns. Therefore, the development of a knowledge-based expert system can help the operators and quality control practitioners to identify the possible sources of variation and take necessary decisive actions.

## 2.4 Shewhart Control Charts

Control charts, also known as Shewhart charts (after Walter A. Shewhart) or process-behavior charts, in statistical process control are tools used to determine if a manufacturing or business process is in a state of statistical control. If analysis of the control chart indicates that the process is currently under control (i.e., is stable, with variation only coming from sources common to the process), then no corrections or changes to process control parameters are needed or desired. In addition, data from the process can be used to predict the future performance of the process. If the chart indicates that the monitored process is not in control, analysis of the chart can help determine the sources of variation, as this will result in degraded process performance (Haridy, S., Wu, Z. (2009)). A process that is stable but operating outside of desired (specification) limits (e.g., scrap rates may be in statistical control but above desired limits) needs to be improved through a deliberate effort to understand the causes of current performance and fundamentally improve the process (Wen & Dwayne 1994).

The control chart is one of the seven basic tools of quality control (Haykin, 1999). Typically control charts are used for time-series data, though they can be used for data that have logical comparability (i.e. you want to compare samples that were taken all at the same time, or the performance of different individuals), however the type of chart used to do this requires consideration (Montgomery,Douglas, C, 2009).

A control chart is a graphical and analytic tool for monitoring process variation. The natural variation in a process can be quantified using a set of control limits. Control limits help distinguish common-cause variation from special-cause variation. Typically, action is taken to eliminate special-cause variation and bring the process back in control. It is also important to quantify the common-cause variation in a process, as this determines process capability.

The Control Chart platform provides a variety of control charts, as well as run charts. To support process improvement initiatives, most of the control chart options display separate control charts for different phases of a project on the same chart, as shown in Table 2.1.

Table 2.1: Types of charts(Montgomery,Douglas, C 2009).

| Chart | Process observation | Process observations relationships | Process observations type | Size of shift to detect |
|---|---|---|---|---|
| $\bar{x}$ and R chart | Quality characteristic measurement within one subgroup | Independent | Variables | Large (≥ 1.5σ) |
| $\bar{x}$ and s chart | Quality characteristic measurement within one subgroup | Independent | Variables | Large (≥ 1.5σ) |
| Shewhart individuals control chart (ImR chart or XmR chart) | Quality characteristic measurement for one observation | Independent | Variables[†] | Large (≥ 1.5σ) |
| Three-way chart | Quality characteristic measurement within one subgroup | Independent | Variables | Large (≥ 1.5σ) |
| p-chart | Fraction nonconforming within one subgroup | Independent | Attributes[†] | Large (≥ 1.5σ) |
| np-chart | Number nonconforming within one subgroup | Independent | Attributes[†] | Large (≥ 1.5σ) |
| c-chart | Number of nonconformances within one subgroup | Independent | Attributes[†] | Large (≥ 1.5σ) |
| u-chart | Nonconformances per unit within one subgroup | Independent | Attributes[†] | Large (≥ 1.5σ) |
| EWMA chart | Exponentially weighted moving average of quality characteristic measurement within one subgroup | Independent | Attributes or variables | Small (< 1.5σ) |
| CUSUM chart | Cumulative sum of quality characteristic measurement within one subgroup | Independent | Attributes or variables | Small (< 1.5σ) |
| Time series model | Quality characteristic measurement within one subgroup | Autocorrelated | Attributes or variables | N/A |
| Regression control chart | Quality characteristic measurement within one subgroup | Dependent of process control variables | Variables | Large (≥ 1.5σ) |

[*]Some practitioners also recommend the use of Individuals charts for attribute data, particularly when the assumptions of either binomially distributed data (p- and np-charts) or Poisson-distributed data (u- and c-charts) are violated (Wen & Dwayne 1994). Two primary justifications are given for this practice. First, normality is not necessary for statistical control, so the Individuals chart may be used with non-normal data (Haykin, 1999). Second, attribute charts derive the measure of dispersion directly from the mean proportion (by assuming a probability distribution), while Individuals charts derive the measure of dispersion from the data, independent of the mean, making Individuals charts more robust than attributes charts to violations of the assumptions about the distribution of the underlying population (Wen & Dwayne 1994). It is sometimes noted that the substitution of the Individuals chart works best for large counts, when the binomial and Poisson distributions approximate a normal distribution. i.e. when the number of trials $n > 1000$ for p- and np-charts or $\lambda > 500$ for u- and c-charts.

The most common use method in current industries is control chart or Shewhart Charts. These control charts are constructed by plotting product's quality variable over time in sequence plot as shown in Figure 2.4.

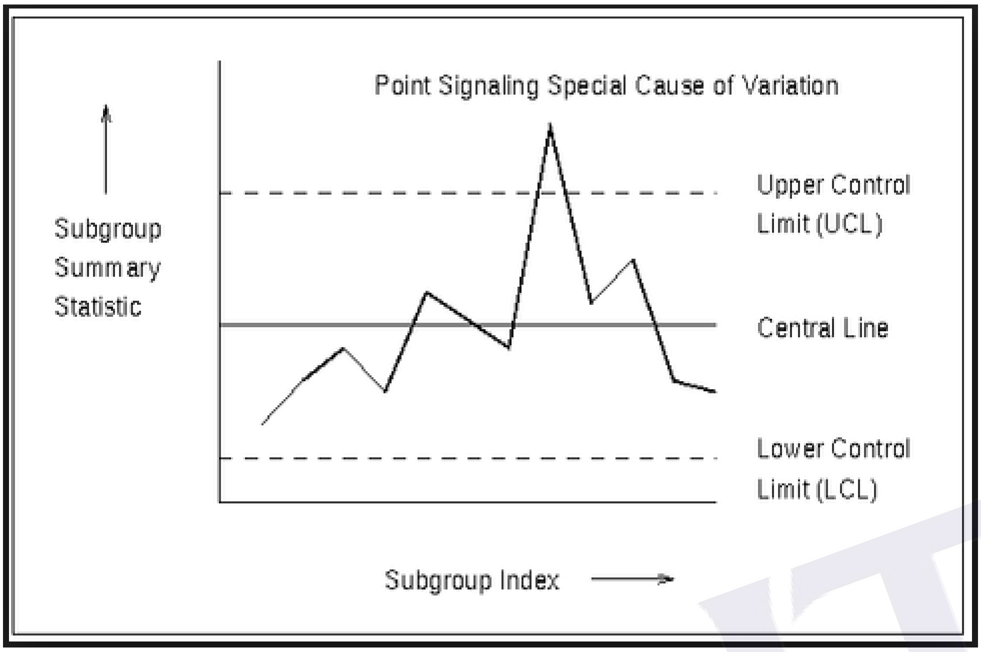



Figure 2.4: Shewhart Charts(Montgomery,Douglas, C 2009).

A control chart contains a center line, an upper control limit and a lower control limit. A point that plots within the control limits indicates the process is in control. In this condition no action is necessary. A point that plots outside the control limits is evidence that the process is out of control. In this condition, investigation and corrective action are required to find and eliminate assignable cause(s) (Umit and Cigdem, 2001). Let $w$ be a sample statistic that measure some quality characteristic of interest and suppose that the mean of $w$ is $\mu w$ and the standard deviation of $w$ is $\sigma w$. Then the center line, upper control limit and lower control limit as shows in equation (2.1).

$$\begin{aligned} \text{UCL} &= \mu w + L\sigma w \\ \text{Center Line} &= \mu w \qquad (2.1) \\ \text{UCL} &= \mu w - L\sigma w \end{aligned}$$

## 2.5 Control Limits

A point falling within the control limits means it fails to reject the null hypothesis that the process is statistically in-control, and a point falling outside the control limits means it rejects the null hypothesis that the process is statistically in-control.

Therefore, the statistical Type I error α (Rejecting the null hypothesis $H_0$ when it is true) applied in Shewhart control chart means the process is concluded as out-of control when it is truly in-control. Same analog, the statistical Type II error β (failing to reject the null hypothesis when it is false) means the process is concluded as in-control when it is truly false.

## 2.6 Average Run Length

The performance of control charts can also be characterized by their average run length. Average run length is the average number of points that must be plotted before a point indicates an out-of-control condition (Montgomery, 1985). We can calculate the average run length for any Shewhart control chart according to:

$$ARL = \frac{1}{P} \tag{2.2}$$

Where P or Type I error is the probability that an out-of-control event occurs. Hence, a control chart with 3 sigma control limits, the average run length will be

$$ARL = \frac{1}{P} = \frac{1}{0.027} = 370 \tag{2.3}$$

This means that if the process remains in-control, in average, there will be one false alarm every 370 samples.

## 2.7 Pattern Recognition in SPC

Pattern recognition is the science of making inferences from perceptual data, using tools from statistics, probability, computational geometry, machine learning, signal processing, and algorithm design (Masood & Hassan, 2010). The techniques of pattern recognition have been successfully used in many areas such as applications in engineering, science, medicine, and business. In particular, advances made during the last half century, now allow computers to interact more effectively with humans and the natural world examples such as speech recognition, word recognition and finger print identification (Wen & Dwayne 1994).

The effectiveness of the use of SPC control charts depends largely on recognizing out-of control conditions in terms of patterns . Guh & Tannock (1999) stated, pattern recognition is an important issue in SPC, as unnatural patterns exhibited by control charts can be associated with specific assignable causes adversely affecting the process. Traditional Shewhart control charts signal only a simple decision, such as within or outside the control limits, based on the most recent observation (Wen & Dwayne, 1994). Control chart pattern recognition (CCPR) has become an active area of research since late 1980s (Masood & Hassan, 2010). Today, control chart pattern recognition has become an active area of research. Zorriassatine, Tannock & O'Brian (2003) provided a useful review on the application for CCPR. However, it is still limited research and updated review on ANN-based CCPR schemes.

There were several pattern recognition approaches done by several researchers. Swift (1987), done a research on SPC control chart pattern recognition using a dichotomous decision tree approach. Swift & Mize (1995) and Cheng (1995), used of expert systems. Expert system also known as rule-based that contain information explicitly. If required, the rules can be modified and updated easily. While the performance of this system was promising, it was reported that the template-matching is currently computationally too expensive to implement in a real-time application scheme (Cheng, 1997).

## 2.8 Artificial Neural Network in Pattern Recognition

Traditionally, statistical process control (SPC) was used only for monitoring and identifying process variation. Advances in SPC charting have moved from merely statistical and economic control to diagnosis purposes through control chart pattern identification. The development in soft computing technology such as artificial intelligence (AI) has encouraged investigation on the application of expert systems, artificial neural network (ANN) and fuzzy sets theory for automated recognition of control chart patterns (CCPs). Application of ANN-based models, among others, has realized the computerized decision making in SPC towards replacing human interpretation. The modernization of the SPC schemes is ultimately aims to diagnose the source of variation with minimum human intervention (Masood & Hassan, 2009).

Recently, many studies used ANNs in order to detect patterns more efficiently than the traditional approach and their goal is the automatic monitoring and diagnosis of the patterns such as shown in Figure 2.5 below.

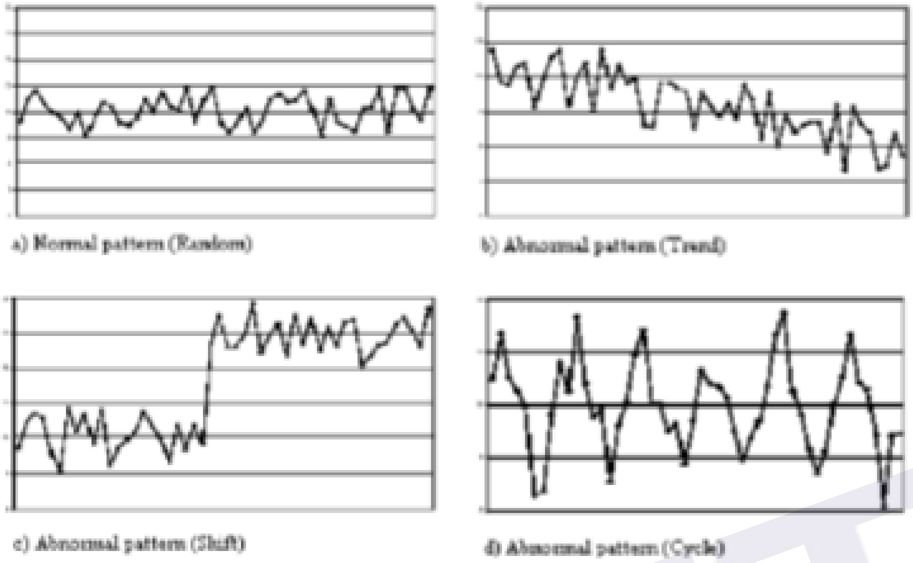



Figure 2.5: Typical normal and abnormal patterns (El-Midany *et* al, 2010)

El-Midany et al (2010) documented, two approaches in applying ANNs to control charts analysis, they are generally:

(i) Uses of ANN to detect deviation in mean and/or variance.

(ii) Uses of ANN to identify abnormal patterns using trained recognizer.

Since late 1980s, control chart pattern recognition (CCPR) has become an active area of research. A useful review on the application of ANN for CCPR was provided by (Zorriassatine & Tannock, 1998). Since then much progress has been made in which the performance of ANN-based CCPR schemes have been enhanced through feature-based and wavelet-denoise input representation techniques, modular and integrated recognizer designs, and multivariate process monitoring and diagnosis. However, there is a lack of updated critical review on such issues.

ANN is a massively parallel-distributed processor that has the ability to learn, recall and generalize knowledge (Haykin, 1999). It is recognized as an important and emerging methodology in the area of classification. ANN is flexible, adaptive and can better handle noise and changes in the patterns. The advantage with an ANN-based pattern recognizer is that it does not require the provision of explicit rules or templates. Rather, it learns to recognize patterns from examples during the training phase. It has the ability to classify an arbitrary pattern not previously encountered. ANN offers useful properties and capabilities such as non-linearity, input and output mapping, adaptability and fault tolerance, among others. These attributes are needed for recognizing and classifying data which are often contaminated with noise, unknown distribution and incomplete as found in CCPs (Schalkoff, 1997; Haykin, 1999).

ANN acquires knowledge through a learning process and inter-neuron connection strengths (synapse weights) are used to store the knowledge. A learning algorithm is used to modify the synapse weights so as to achieve the target. ANN can tailor itself to the training data. A well-trained ANN is able to generalize knowledge. It will produce a reasonable output for input that has never been encountered during training/learning. Although ANN training requires considerable computation, the recall process is very fast. ANN is also suitable for implementation using very-large-scale-integrated (VLSI) technology such as in the form of chip that can replace the need for continuously monitoring by personal computer (Scalkoff, 1997; Haykin, 1999).

# CHAPTER 3

# METHODOLOGY

## 3.1 Introduction

Methodology is very important to determine the direction, guidelines and methods in performing the project. This chapter explains the working procedures to complete the whole project. Methodology also indicates the procedural steps that need to be follow to ensure that the project will be completed on time. All procedures and methods have been listed down to give a guideline on project progress and to ensure the project completed as planned. All the activities were listed on the Gantt chart as shown in Figure A.1 (in Appendix A). Figure 3.1 illustrated the explanation on the methodologies in integrated-ANN scheme in monitoring and diagnosis of bivariate process variation in mean shifts.

In this research, two-stage monitoring scheme for improving the balanced monitoring and accurate diagnosis was investigated by Integrated SPC-ANN. Framework for the proposed scheme are summarized in Figure 3.2.

## 3.2 Project Methodology

Overview of project methodology for this research are showed as below:-

### 3.2.1 Research Overview

(i) Understand the concept and the application of SPC.

(ii) Understand the merit of SPC in process monitoring and diagnosing.

(iii) Understand various type of control chart.

### 3.2.2 Data Generator

To generate the data for ANN Trainer by using a scheme based on the input representation of the simulation data for training and testing validation testing.

### 3.2.3 ANN Trainer

Setup and run ANN trainer to train:-

(i) Features-Based ANN recognizer.

(ii) Raw Data-Based ANN recognizer.

(iii) Raw data and improved set of statistical features were applied separately into training for improving pattern discrimination capability.

(iv) Run until achieve the targets that have been set so as to be the scheme in a well-trained and record the data from.

### 3.2.4 Control Chart

Bivariate :

(i) MEWMA;

(ii) Synergistic MEWMA-ANN scheme

### 3.2.5 Expected Results

(i) Select bivariate variables with low, moderate & high correlation.

(ii) Interpret scatter plot diagram pattern.

(iii) Performance investigation of control charts in term of monitoring and diagnosis by using Integrated SPC – ANN.

(iv) Project conclusions.

The research methodology flow chart of research that was carried out throughout a year is demonstrates in Fig. 3.1 below.

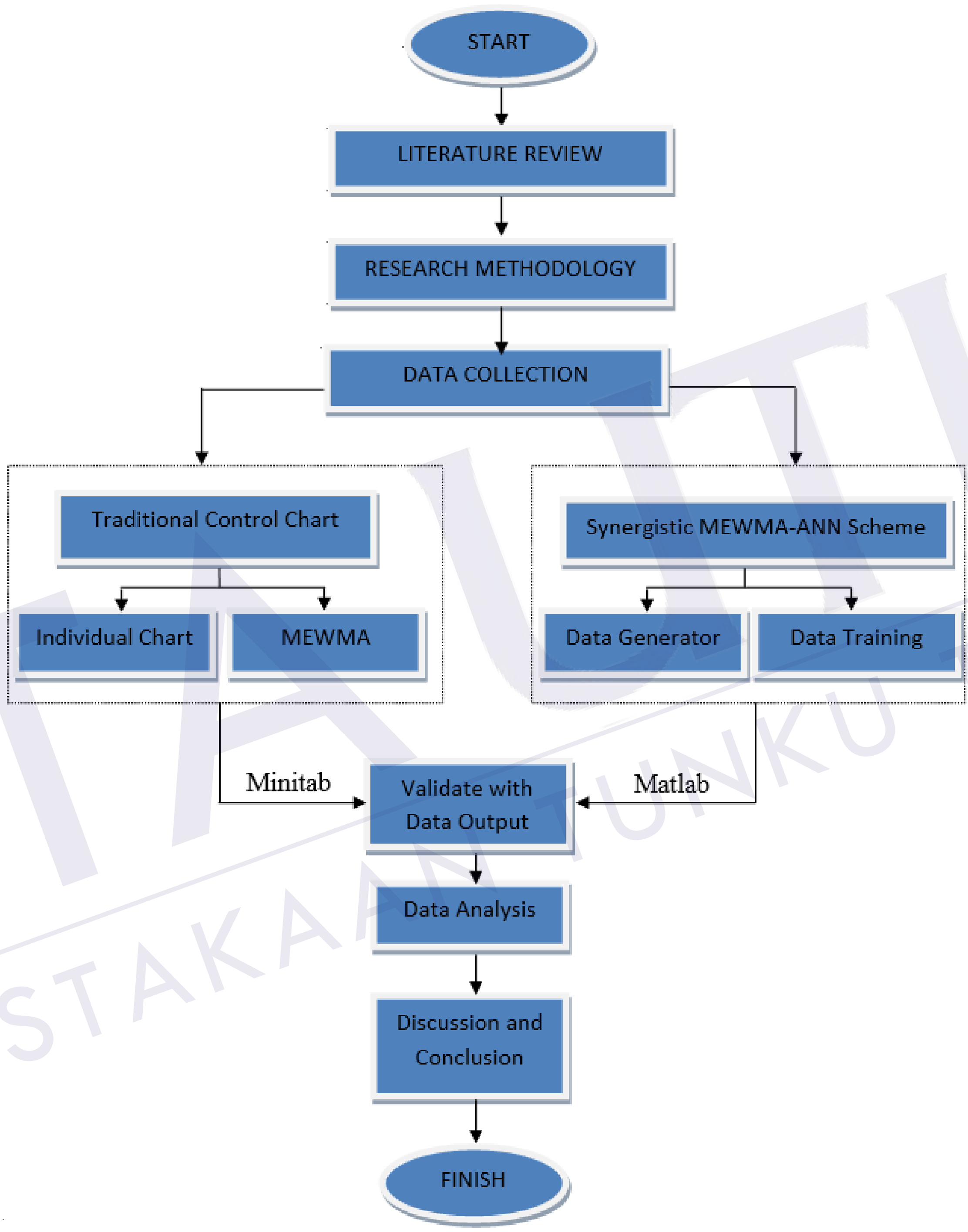



Figure 3.1: Research flow chart

### 3.3 An integrated MEWMA-ANN scheme

An integrated MEWMA-ANN scheme was develop based on two-stages monitoring and diagnosis approach as shown in Figure 3.2. Process monitoring refers to the identification of process status either in a statistically stable or unstable state, whereas process diagnosis refers to the identification of the source variables of an unstable process. In the first stage monitoring, the MEWMA control chart is used for triggering mean shifts based on ‘one point out-of-control’. Once the mean shift is triggered, the Synergistic-ANN recognizer is then used to perform second stage monitoring and diagnosis by recognizing data stream pattern contained points out-of-control as truly unstable or not.

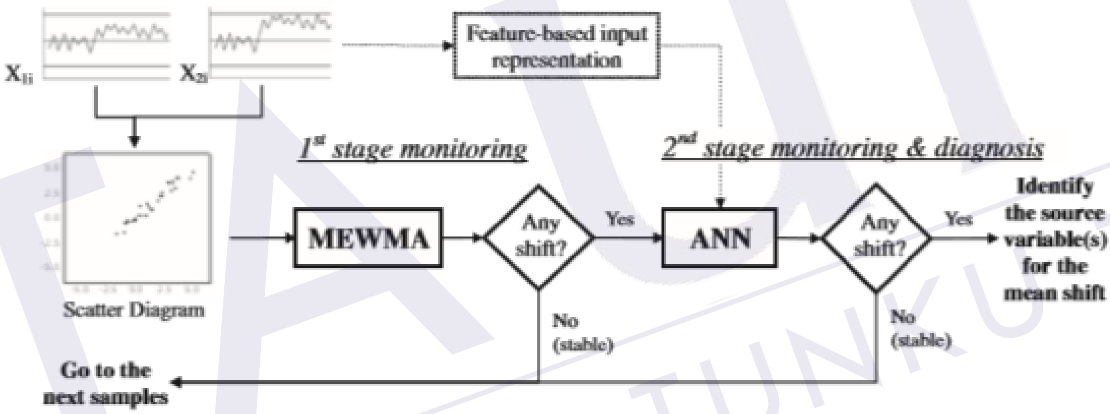



Figure 3.2: Conceptual diagram of an Integrated MEWMA-ANN scheme (Masood, 2014)

### 3.4 Modeling of Process Pattern

A large number of multivariate collated samples is required to perform a through pattern recognition study through training and testing the ANN recognizers. Ideally such samples should be tapped from real process environment. Since they are not economically available, there is a need for modeling of process patterns for synthetically generating analysis data.

Consider a bivariate correlated process. Let $X_{1i} = (X_{1\text{-}1},\ldots,X_{1\text{-}24})$ and $X_{2i} = (X_{2\text{-}1},\ldots,X_{2\text{-}24})$ represent data streams for process variable process 1 and variable process 2 based on observation window of 24 samples. Observation windows for both variables start with samples $i^{th} = (1,\ldots 24)$. Then, it is followed with $(i^{th} + 1)$, $(i^{th} + 2)$,…., and so on.

The occurrence of assignable causes over $X_{1i}$ and/or $X_{2i}$ can be identified by common causable pattern such as sudden shifts, trends, cyclic, systematic and mixture. It involves seven possible conditions for multivariate process mean shifts with positive cross correlation ($\rho$) as followed:-

Normal (0, 0); Both X1i and X2i are stable

Up –shift (1, 0); X1i in upward shift and X2i remain stable

Up –shift (0, 1); X2i in upward shift and X1i remain stable

Up –shift (1, 1); Both X1i and X2i in upward shift

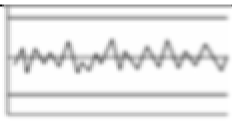
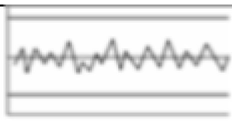
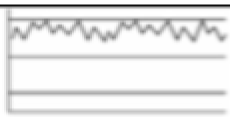
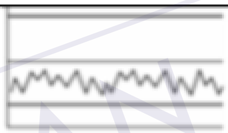
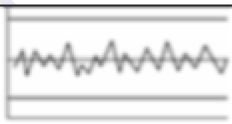
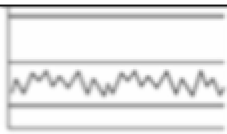
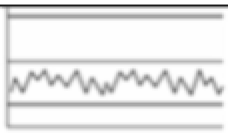
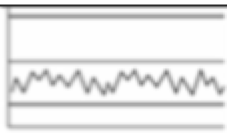
Down –shift (-1, 0); X1i in downward shift and X2i is stable

Down –shift (0, -1); X2i in downward shift and X1i is stable

Down –shift (-1, -1); Both X1i and X2i in downward shift

Seven possible conditions of multivariate process mean shifts with positive cross correlation (ρ) were involved in this project as summarized in Table 3.1.

Table 3.1: Conditions of multivariate process mean shifts

| Pattern Conditions | $X_{1i}$ | $X_{2i}$ |
|---|---|---|
| Normal (0, 0) |  Normal |  Normal |
| Up –shift (1, 0) |  Upward |  Normal |
| Up –shift (0, 1) |  Normal |  Upward |
| Up –shift (1, 1) |  Upward |  Upward |
| Down –shift (-1, 0) |  Downward |  Normal |
| Down –shift (0, -1) |  Normal |  Downward |
| Down –shift (-1, -1) |  Downward |  Downward |

## REFERENCES


Brunzell, H., & Eriksson, J. (2000). Feature reduction for classification of multi dimensional data. *Pattern Recognition*, 33(12), 1741–1748.

Chen, L.H. & Wang, T.Y. (2004). Artificial Neural Networks to Classify Mean Shifts from Multivariate X2 Chart Signals. Computers & Industrial Engineering. 47(2004), pp. 195-205.

Chen, Z., Lu, S., Lam, S. (2007). A Hybrid System for SPC Concurrent Pattern Recognition. Advanced Engineering Informatics. 21(2007), pp. 303-310.

Champ, C.W. and Woodall, W.H., "Exact result for Shewhart control chart with supplementary runs rules", *Technometrics*, Vol. 29 No. 4, pp. 393-9, November 1987.

Cheng CS (1995) A multi-layer neural network model for detecting changes in the process mean. Comput IND Eng 28:51-61.

Cheng, C.S., 1997. A neural network approach for the analysis of control chart patterns. *International Journal of Production Research* 35, 667–697.

Cheng, H.P., Cheng, C.S. (2009). Control Chart Pattern Recognition Using Wavelet Analysis and Neural Networks. Journal of Quality, 16(25), 311-321.

Electric W, (1958). Statistical Quality Control handbook. Western Electric Company, I ndianapolis.

El-Midany, T.T., El-Baz, M.A., Abd-Elwahed, M.S. (2010). A Proposed Framework for Control Chart Pattern Recognition in Multivariate Process Using Artificial Neural Networks. *Expert System with Applications. 37(2010),* pp. 1035-1042.

Haykin S (1999) Neural networks: a comprehensive foundation. Prentice Hall, New Jersey.

Haridy, S., Wu, Z. (2009). Univariate and Multivariate Control Charts for Monitoring Dynamic-Behavior Process: A Case Study. *Journal of Industrial Engineering and Management, 2(3).pp. 464-498.*

Hassan, A., Shariff Nabi Baksh, M., Shaharoun, A.M., Jamaluddin, H. (1998). Improved SPC Chart Pattern Recognition Using Statistical Features. *International Journal of Production Research, 41(7).pp. 1587-1603.*

Guh RS, Tannock JDT, O'Brien C (1999) Intellispc: a hybrid intelligence tool for on-line economical statistical process control. Expert Syst Appl 17:195-212.

Hwarng, H.B.(2008). Toward Identifying The Source of Mean Shifts in Multivariate SPC: A Neural Network Approach. *International Journal of Production Research. 46(20), pp. 5531-5559.*

Klosgen, W., & Zytkow, J. M. (2002). *Handbook of data mining and knowledge discovery*. London: Oxford University Press.

Kourti, T., & MacGregor, J. F. (1996). Multivariate SPC methods for process and product monitoring. *Journal of Quality Technology*, 28(4), 409–428.

Lucas, J. M., & Saccucci, M. S. (1990). Exponentially weighted moving average control schemes: Properties and enhancements. *Technometrics*, 32(1), 1–12.

Masood, I. & Hassan, A. (2009). Synergistic-ANN Recognizers for Monitoring and Diagnosis of Multivariate Process Shift Patterns. *International Journal of Soft Computing and Pattern Recognition*, 2 (9), pp. 336-355.

Masood, I. & Hassan, A. (2010). Improved Recognition of Multivariate Process Mean Shift Patterns Using Synergistic-ANN Recognizer. *The 11th Asia Pacific Industrial Engineering and Management Systems Conference.7-10 December 2010: APIFM.2010.*

Masood, I. & Hassan, A. (2010). Issues in Development of Artificial Neural Network-Based Control Chart Pattern Recognition Schemes. *European Journal of Scientific Research, 39(3), pp. 336-355.*

Masood, I. & Hassan, A. (2012). Pattern Recognition for Bivariate Process Mean Shifts Using Feature-Based Artificial Neural Network. *The International Journal of Advanced Manufacturing Technology, 24(9&10).*

Masood, I. & Hassan, A. (2013). A Framework for Multivariate Process Monitoring and Diagnosis. *Applied Mechanics and Materials, 315(2013), pp. 374-379.*

Masood, I. & Hassan, A. (2013). Multivariate Process Monitoring and Diagnosis: A Case Study. *Applied Mechanics and Materials, 315(2013), pp. 606-611.*

Masood, I. & Hassan, A. (2014). *Journal of Bivariate quality control using two-stage intelligent monitoring scheme*. 41 (2014) 7579-7595.

Montgomery, D.C (2005) Introduction to Statistical Quality Control. Wiley, New Jersey.

Montgomery, D.C (2009). Introduction to Statistical Quality Control. 6th ed. USA. John Wiley & Sons, Inc. New Jersey.

Montgomery, D.C. (2013). Statistical Quality Control – A Modern Introduction. 7th ed. USA. John Wiley & Sons, Inc. New Jersey.

Montgomery, D.C., *Introduction to Statistical Quality Control*, John Wiley & Sons, New York, NY, 1985.

Oguz, H., Saritas, I., Badyan, H.E, (2010). Prediction of Diesel Engine Performance Using Biodeiesel with Artifical Neural Network. 37, 6579–6586.

Lowry, C. A., & Montgomery, D. C. (1995). A review of multivariate control charts. IIE Transactions, 27(6), 800–810.

Page, E.S., “Continuous Inspection Schemes”, *Biometrika*, Vol. 41, 1954, pp. 100-14.

Roberts, S.W., “Control Charts Tests Based on Geometric Moving Averages”, *Technometrics*, Vol. 1, 1959, pp. 239-50.

Schalkoff, R.J., 1997. “Artificial neural networks”, *McGraw-Hill*, New York.

Swift, J. A., & Mize, J. H. (1995). Out-of-control pattern recognition and analysis for quality control charts using LISP-based systems. *Computers and Industrial Engineering, 28(1), 81–91.*

Swift, J.A. (1987). Development of A Knowledge Based Expert System for Control Chart Pattern Recognition and Analysis, *unpublished doctoral dissertation*, The Oklahoma State University.

Umit, F. & Cigdem, A. (2001), Multivariate Quality Control: A Historical Perspective pp. 54-65.

Wen, H.C. & Dwayne, A.R. (1994). A Methodology of Pattern Recognition Schemes for Two Variables in SPC.*International Journal of Quality & Reliability Management. 12(3)*, pp. 86-107.

Yu, J.B. & Xi, L.F. (2009). Artificial Neural Networks to Classify Mean Shifts from Multivariate X2 Chart Signals. *Expert System with Applications. 36(2009),* pp. 909-9221.

Zainal Abidin, N.Z. & Masood, I. (2012). *Design of an Artificial Neural Network Pattern Recognition Scheme Using Full Factorial Experiment.* Bachelor Degree’s Thesis, University Tun Hussein Onn Malaysia.

Zorriassatine F, Tannock JDT (1998) *A review of neural networks for statistical process control*. J Intell Manuf 9:209-224

Zorriassatine, F., Tannock, J.D.T., and O’Brien, C, (2003). Using Novelty Detection to Identify Abnormilities Caused by Mean Shifts in Bivariate Processes. *Computers and Industrial Engineering, 44(2003),* pp. 385-408.